ELSEVIER

Bacteriophage-based assays for the rapid detection of rifampicin resistance in *Mycobacterium tuberculosis*: a meta-analysis

Madhukar Pai[a,b,*], **Shriprakash Kalantri**[a,c], **Lisa Pascopella**[d], **Lee W. Riley**[a], **Arthur L. Reingold**[a]

[a]*Division of Epidemiology, School of Public Health, University of California, 140, Warren Hall, Berkeley, CA 94720, USA*
[b]*Division of Pulmonary and Critical Care Medicine, San Francisco General Hospital, University of California, San Francisco, CA, USA*
[c]*Department of Medicine, Mahatma Gandhi Institute of Medical Sciences, Sevagram, India*
[d]*Surveillance and Epidemiology Section, Tuberculosis Control Branch, California Department of Health Services, Berkeley, CA, USA*

Abstract *Objective*: To summarize, using meta-analysis, the accuracy of bacteriophage-based assays for the detection of rifampicin resistance in *Mycobacterium tuberculosis*.

Methods: By searching multiple databases and sources we identified a total of 21 studies eligible for meta-analysis. Of these, 14 studies used phage amplification assays (including eight studies on the commercial FASTPlaque-TB® kits), and seven used luciferase reporter phage (LRP) assays. Sensitivity, specificity, and agreement between phage assay and reference standard (e.g. agar proportion method or BACTEC 460) results were the main outcomes of interest.

Results: When performed on culture isolates ($N = 19$ studies), phage assays appear to have relatively high sensitivity and specificity. Eleven of 19 (58%) studies reported sensitivity and specificity estimates $\geq 95\%$, and 13 of 19 (68%) studies reported $\geq 95\%$ agreement with reference standard results. Specificity estimates were slightly lower and more variable than sensitivity; 5 of 19 (26%) studies reported specificity $< 90\%$. Only two studies performed phage assays directly on sputum specimens; although one study reported sensitivity and specificity of 100 and 99%, respectively, another reported sensitivity of 86% and specificity of 73%.

Conclusions: Current evidence is largely restricted to the use of phage assays for the detection of rifampicin resistance in culture isolates. When used on culture

* Corresponding author. Address: Division of Epidemiology, School of Public Health, University of California, 140, Warren Hall, Berkeley, CA 94720, USA. Tel.: +1 510 388 7137; fax: +1 510 643 4927.
 E-mail address: madhupai@berkeley.edu (M. Pai).

isolates, these assays appear to have high sensitivity, but variable and slightly lower specificity. In contrast, evidence is lacking on the accuracy of these assays when they are directly applied to sputum specimens. If phage-based assays can be directly used on clinical specimens and if they are shown to have high accuracy, they have the potential to improve the diagnosis of MDR-TB. However, before phage assays can be successfully used in routine practice, several concerns have to be addressed, including unexplained false positives in some studies, potential for contamination and indeterminate results.

Introduction

Multidrug-resistant tuberculosis (MDR-TB), defined as resistance to at least isoniazid (INH) and rifampicin (RIF), is a major threat to tuberculosis control.[1] According to the Global Project on anti-tuberculosis drug resistance surveillance (2004), drug-resistant TB, including MDR-TB, is found in all regions of the world.[2] Globally, among new TB patients, the median prevalence was 10.2% for any drug resistance, and 1.1% for MDR-TB.[2] Among previously treated TB patients, the median prevalence was 18.4% for any drug resistance, and 7% for MDR-TB.[2] This project also showed significant increases in the prevalence of drug resistance in a number of settings between 1994 and 2002.[2]

Early detection of MDR-TB and prompt treatment with appropriate regimens including second-line drugs can reduce morbidity and mortality in patients with drug-resistant TB, and also reduce transmission of drug-resistant bacilli.[1] However, the diagnosis of drug-resistant TB poses several challenges.[3,4] Conventional drug susceptibility tests (DST) include the absolute concentration method, proportion method, and the resistance ratio method.[3-5] These tests are typically performed on cultures isolated from clinical specimens. They are time-consuming (with solid culture media results are obtained 3–4 weeks after primary isolation), tedious, and not easily accessible in many developing countries.[3-5] Newer techniques (particularly those involving liquid cultures such as the BACTEC®, MB/BacT® and MGIT® systems) are more rapid than conventional agar-based methods. However, they are expensive, require sophisticated laboratory infrastructure and trained personnel, and poorly suited for many laboratories in resource-limited settings.[3-5]

In this context, substantial effort has been directed towards the development of rapid tests for drug resistance, particularly focused on detection of rifampicin resistance.[5-7] Because rifampicin mono-resistance is rare, resistance to rifampicin is considered a surrogate marker of MDR-TB.

However, the correlation between rifampicin mono-resistance and MDR-TB may vary across populations depending on a number of factors. As a result, in some populations rifampicin mono-resistance may not be an accurate marker of MDR-TB.[2] Newer tests for rifampicin resistance include genotypic tests (e.g. nucleic acid amplification (NAA) tests, DNA sequencing (e.g. single nucleotide polymorphism (SNP) analysis), line probe assays, DNA microarrays, and molecular beacons), and phenotypic tests (e.g. mycobacteriophage-based assays).[3,5-7] Molecular genotypic tests for rifampicin resistance aim to detect mutations that are associated with drug resistance.[5-7] More than 95% of rifampicin-resistant *M. tuberculosis* strains contain a mutation in a gene that encodes the b-subunit of RNA polymerase (*rpoB*). Phenotypic tests, in contrast, detect drug resistance regardless of its genetic basis, and compare the growth of *M. tuberculosis* in the presence and absence of the drug being tested. Among the phenotypic tests that have been evaluated, assays based on mycobacteriophages have shown some promise.[5,6,8-12] Phage-based assays have been evaluated for diagnosis of pulmonary TB, as well as detection of drug resistance.[8,9,11] In a previous meta-analysis, we summarized the accuracy of phage-based assays for the detection of *M. tuberculosis* in clinical specimens.[13] In this meta-analysis, we summarize the accuracy of phage-based assays for detection of rifampicin resistance.

Phage-based assays utilize bacteriophages to infect and detect the presence of *M. tuberculosis* in clinical specimens and culture isolates.[8,9] Two main approaches are used to detect *M. tuberculosis*: (1) amplification of phages (Fig. 1) after their infection of *M. tuberculosis*, followed by detection of progeny phages using Sensor cells (plaque formation);[12] and (2) detection of light produced by luciferase reporter phages (LRP) after their infection of live *M. tuberculosis* (Fig. 2).[10] Assays based on phage amplification methods have been more frequently evaluated than tests based on LRP. Also, amplification based assays are commercially
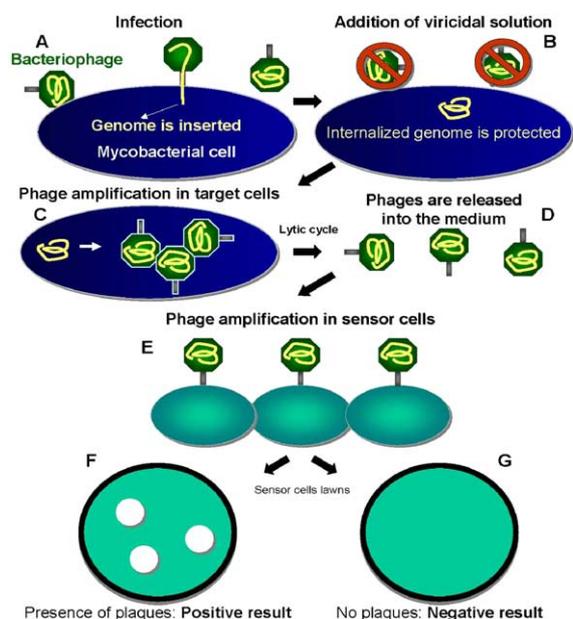
Figure 1 An overview of the phage amplification assay. Source: Hazbon, 2004[6]; reproduced with permission from Biomedica.

available, whereas LRP tests are still under development and restricted to research settings. In both systems, drug resistance is diagnosed when *M. tuberculosis* is detected in samples that contain the drug (e.g. RIF). When phage-based assays fail to detect *M. tuberculosis* in drug containing samples, the strains are classified as drug-sensitive.[8,9]

Phage-based assays are available as commercial kits (e.g. FASTPlaque-TB® and PhageTek MB®, a variant of the FASTPlaque-TB, Biotec Laboratories Ltd, U.K.) and as in-house (laboratory-developed) assays.[8,9] In-house assays use either amplification technology (e.g. PhaB) or LRPs. Some of the phage-based assays are specifically designed to rapidly detect rifampicin resistance (e.g. FASTPlaque-TB-MDRi®, earlier called FASTPlaque-TB-RIF®) in culture isolates. Newer kits are being developed for the rapid detection of drug resistance directly from clinical specimens such as sputum (e.g. FASTPlaque-TB-Response®).
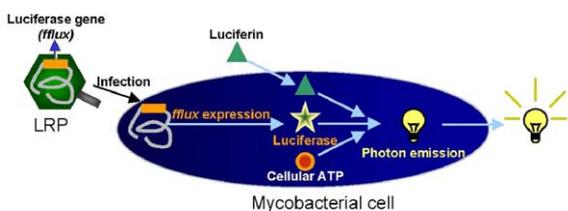
Source: Hazbon 2004[6] (reproduced with permission from *Biomedica*)

Figure 2 An overview of the luciferase reporter phage assay.

To evaluate the accuracy of phage-based assays for rifampicin resistance, we conducted a systematic review and meta-analysis. A secondary objective was to assess if studies produced heterogeneous estimates of accuracy, and to evaluate if such heterogeneity is due to variability in test characteristics and study methodology.

Methods

Our meta-analysis was conducted according to a pre-specified protocol. The outline of the methods used in our systematic review has been described elsewhere.[14,15] Briefly, the review involved the following major steps: (1) formulation of the review question, (2) a comprehensive, systematic search and selection of primary studies, (3) critical appraisal of included studies for quality, (4) data extraction and contact with authors for additional information, (5) synthesis and summary of study results, and (6) interpretation of the results.

Study eligibility

We aimed to include original clinical studies that evaluated the sensitivity and specificity of phage-based assays (either phage amplification or LRP) for the detection of rifampicin-resistant *M. tuberculosis* in clinical specimens or isolates. Both commercial tests and in-house assays were included, but separately analysed. The following studies were excluded: animal experiments, proof of principle studies on development of new assays, review articles, letters, and commentaries, and studies that did not evaluate both sensitivity and specificity in the same population.

With respect to study design, we included diagnostic studies (either case-control or cross-sectional) that evaluated the accuracy (sensitivity and specificity) of phage-based assays against a reference standard. Any of the following conventional drug susceptibility tests were accepted as reference standards: absolute concentration method, proportion method, resistance ratio method, and radiometric BACTEC 460 method.

Search strategy

We searched the following electronic databases for primary studies and conference abstracts: PubMed (1985–2004), Web of Science (1985–2004), Embase (1988–2004), BIOSIS (1993–2004), Cochrane Library (2004), and Google Scholar (December 2004). All searches were up to date as of December 2004.

We did not impose language restrictions in our searches. The keywords and search terms used included 'tuberculosis', '*mycobacterium tuberculosis*', 'mycobacteria', 'bacteriophages', 'myco-bacteriophage', 'phage*', 'Fastplaque', 'phage amplification', 'phage-based', 'bacteriophage-based', 'luciferase', 'sensitivity and specificity', 'accuracy' and 'predictive value'. We also contacted experts in the field to identify ongoing and unpublished studies, and searched the reference lists from the primary studies and review articles. To identify all relevant studies of commercial assays, we contacted and obtained lists of studies from commercial kit manufacturers.

Study selection

Two reviewers (SP and MP) independently screened the citations. Citations that were relevant on the first screen of titles and abstracts were evaluated further by review of full-text reports. Disagreements between the reviewers were resolved by consensus. A list of excluded studies, along with the reasons for exclusion is available from the authors on request.

Data extraction and assessment of study quality

The final set of included articles was assessed by one reviewer (MP), who extracted data from all the studies using a piloted data extraction form. A second reviewer (SP) independently extracted data from a subset (one-fourth) of the included studies. The inter-rater agreement between the two reviewers for sensitivity and specificity data was 100%. Data retrieved from the reports included methodological quality, phage test characteristics, reference standards employed, outcome data (sensitivity, specificity, and agreement between phage results and conventional DST), and proportion of results that were indeterminate, contaminated or had intermediate sensitivity. Since discrepant analysis (where discordant results between index test and reference standard are resolved, post-hoc, using clinical or other laboratory data) may be a potential source of bias in diagnostic evaluations,[16] we preferentially included unresolved data where available. This approach is likely to result in more conservative estimates of accuracy, and more likely to reflect the performance of tests in routine practice rather than research settings.

We assessed the quality of the studies by using the following criteria: selection bias (consecutive or random sampling of patients/specimens versus studies that used neither method), blinding (single/double blind versus unblinded interpretation of phage test and reference standard results), and potential for verification bias (complete versus partial/differential verification of index test results by reference standards). We contacted authors, when necessary, to obtain additional information on study quality and results. Additional information was obtained from the authors for 16 of 21 (76%) studies.

Data synthesis and meta-analysis

To compute sensitivity and specificity of phage-based assays, we cross-tabulated the phage drug susceptibility results against those obtained from the conventional DST (e.g. BACTEC). Sensitivity (true positive rate) refers to the proportion of rifampicin-resistant strains that are correctly identified as resistant by phage-based assays. Specificity (true negative rate) refers to the proportion of rifampicin-susceptible strains that are correctly identified as susceptible by phage-based assays. For the computation of these measures, most studies excluded test results that were uninterpretable, contaminated, indeterminate, or had intermediate drug sensitivity. In addition to sensitivity and specificity, we computed measures of agreement to evaluate the correlation between conventional DST and phage assay results. Simple proportion agreement and kappa estimates were computed for each study, along with 95% confidence intervals (95% CI).

We used standard methods recommended for diagnostic meta-analyses.[15,17,18] Data were analysed using Meta-DiSc (version 1.1.1) software. Our analyses focused on sensitivity (true positive rate [TPR]), and specificity (1-false positive rate [FPR]) for the detection of rifampicin resistance. Because TPR and FPR are correlated and vary with the thresholds (cut-points for determining test positives) employed in the original studies, we did not pool the sensitivity and specificity estimates separately; instead we analysed TPR and FPR as pairs, and explored the effect of variability in cut-points on study results.[15,17,18]

We summarized the joint distribution of TPR and FPR using the summary receiver operating characteristic (SROC) curve.[19] Unlike a traditional ROC plot that explores the effect of varying thresholds on sensitivity and specificity in a single study, each data point in the SROC space represents an individual study. As described by Littenberg and Moses, the SROC curve is obtained by fitting a

regression curve to pairs of TPR and FPR.[19] To facilitate the SROC analysis, a correction was made to account for zero values in any cell of the 2×2 table generated from each study—0.5 was added to all cells in the 2×2 table. The SROC plots, therefore, display the zero cell-corrected data.

The SROC curve and the area under it present an overall summary of test performance, and display the trade off between sensitivity and specificity. A symmetric, shoulder-like SROC curve suggests that variability in thresholds employed could, in part, explain variability in study results.[15,17-19] It also suggests a common, homogeneous underlying diagnostic odds ratio that does not change with the diagnostic threshold. The area under the ROC curve (AUC) and the Q* index are useful global summaries of the SROC curve and test performance.[19,20] An area under the ROC curve of 1.0 indicates perfect discriminatory ability.[19,20] The Q* index, defined by the point where sensitivity equals specificity on the SROC curve, is that point on the SROC curve that is intersected by the anti-diagonal. A Q* value of 1.0 indicates 100% accuracy (sensitivity and specificity of 1.0).[19,20] Thus, the closer the Q* value is to 1.0, the more accurate the test.

In meta-analyses, heterogeneity refers to a substantial degree of variability in study results.[21] Such heterogeneity could be due to variability in thresholds, disease spectrum, assay methods, and study quality across studies.[21] In the presence of significant heterogeneity, pooled, summary estimates from meta-analyses are hard to interpret. We investigated heterogeneity using subgroup (stratified) analyses. To account for the major differences in assay methods, we analysed phage amplification assays separately from LRP tests. Within phage amplification assays, commercial assays were evaluated separately from in-house assays. In addition, tests performed on culture isolates were evaluated separately from assays performed on clinical specimens. These subgroups were identified a priori and pre-specified in the study protocol. Subgroup analyses also allowed us to determine if any particular phage-based method was more likely to be associated with higher accuracy.

Results

We identified a total of 540 unique citations from all literature searches. Of these, a total of 21 articles[12,22-41] (17 English, 1 Chinese, and 1 Turkish), including two conference abstracts[22,41] (English) were eligible for inclusion in our meta-analysis. However, one study was excluded from the analysis because it had only a single case of confirmed rifampicin resistance.[28] One article reported outcome data separately for two variants of the luciferase reporter phage assay (Bronx box versus luminometric detection methods).[32] These were counted as separate studies. Our final analyses, therefore, had 21 studies.

Description of included studies

The Table 1 presents the characteristics of the 21 included studies, along with data on methodological quality and outcomes. Of these studies, 14 studies[12,22-25,29-31,33,34,36-38,40] evaluated amplification-based assays, whereas seven studies[26,27,32,35,39,41] evaluated LRP assays. Among the 14 phage amplification assays, eight studies[22-25,29,33,34,38] used commercial assays (FASTPlaque-TB), and six studies used in-house assays.[12,30,31,36,37,40] The average sample size was 85 specimens or isolates (range 22–201). On average, studies included fewer resistant than sensitive strains (mean 27 versus 58, respectively).

BACTEC 460 and proportion methods were the most commonly employed reference tests for conventional DST. Eight of 21 (38%) studies recruited patients or specimens using a random or consecutive sampling strategy. Twelve of 21 (57%) studies used a study design that included blinded interpretation of either the phage test or the reference standard results (single or double blinded). None of the studies had potential for verification bias.

With respect to assay characteristics, most studies used culture isolates rather than clinical specimens for the phage assays. Also studies employed varying definitions of drug resistance. For example, a majority of the studies on the FASTPlaque-TB-RIF assay used ≥50 plaques on the RIF—containing plate for classifying a strain as resistant. One study used a cut-point of ≥20 plaques. Cut-points used in LRP studies also varied, depending on the method used for detection of luciferase activity (luminometry versus photographic methods). The proportion of phage results that were indeterminate, and/or contaminated or had intermediate sensitivity, at least on initial testing, varied from 0 to 17%.

Accuracy of phage amplification assays

Of the 14 studies that evaluated phage amplification assays, commercial kits were used in eight studies. These kits included the FASTPlaque-TB-RIF (now called FASTPlaque-TB-MDRi), and FASTPlaque-TB-Response assays. The remaining six studies

Table 1 Description of studies in the meta-analysis and measures of test accuracy

Author	Country	Phage test	Reference standard DST	Sample	Consecutive or random selection of specimens or patients	Single or double blinded	Sample size (number of resistant strains/sensitive strains)	% indeterminate, contaminated or intermediate sensitivity on initial testing	Sensitivity (95% CI)	Specificity (95% CI)	Agreement (Kappa)
Phage amplification assays (Commercial)											
Performed directly on sputum specimens											
Butt (2004)[29]	Pakistan	FASTPlaque-TB-RIF	BACTEC 460	Sputum	Yes	Yes	28/11	3	0.86 (0.67, 0.96)	0.73 (0.39, 0.94)	82% (0.57)
Albert (2004)[24]	South Africa	FASTPlaque-TB-Response	Proportion method	Sputum	Yes	Yes	10/135	17	1.0 (0.69, 1.0)	0.99 (0.96, 1.0)	99% (0.95)
Performed on culture isolates											
Kisa (2003)[33]	Turkey	FASTPlaque-TB-RIF	BACTEC 460	Isolate	NR	NR	21/67	3	1.0 (0.84, 1.0)	0.93 (0.83, 0.97)	94% (0.86)
Albert (2002)[25]	South Africa	FASTPlaque-TB-RIF	BACTEC 460	Isolate	No	Yes	42/91	2	1.0 (0.92, 1.0)	0.98 (0.39, 0.94)	98% (0.97)
Krishnamurthy (2002)[34]	India	FASTPlaque-TB-RIF	Proportion method	Isolate	No	No	73/12	0	0.96 (0.88, 0.99)	1.0 (0.74, 1.0)	96% (0.87)
Albert (2001)[23]	South Africa	FASTPlaque-TB-RIF	Proportion method	Isolate	No	No	76/107	2	1.0 (0.95, 1.0)	0.98 (0.93, 0.99)	99% (0.98)
Oguz (2002)[38]	Turkey	FASTPlaque-TB-RIF	Proportion method	Isolate	NR	NR	21/11	0	0.81 (0.58, 0.95)	0.82 (0.48, 0.98)	81% (0.60)
Aktepe (2001)[22]	Turkey	FASTPlaque-TB-RIF	Proportion method	Isolate	NR	NR	9/31	0	1.0 (0.84, 1.0)	0.74 (0.55, 0.88)	80% (0.56)
Phage amplification assays (in-house) performed on culture isolates											
Gali (2003)[31]	Spain	In-house (D29)	BACTEC 460	Isolate	No	Yes	18/71	0	1.0 (0.82, 1.0)	1.0 (0.95, 1.0)	100% (1.0)
Mani (2003)[36]	India	In-house (D29)	Absolute concentration method	Isolate	No	Yes	101/100	0	0.97 (0.92, 0.99)	0.84 (0.75, 1.0)	91% (0.91)
McNerney (2000)[37]	South Africa, U.K.	In-house (D29)	Proportion method + BACTEC	Isolate	No	Yes	17/20	0	1.0 (0.81, 1.0)	1.0 (0.83, 1.0)	100% (1.0)
Eltringham (1999)[30]	U.K.	In-house (D29)	Resistance ratio method	Isolate	NR	NR	31/46	15	1.0 (0.89, 1.0)	1.0 (0.92, 1.0)	100% (1.0)
Simboli (2005)[40]	Argentina	In-house (D29)	Proportion method	Isolate	Yes	Yes	42/87	7	1.0 (0.92, 1.0)	0.99 (0.94, 1.0)	99% (0.98)
Wilson (1997)[12]	U.K.	In-house (D29)	Absolute concentration+ resistance ratio	Isolate	NR	NR	9/37	0	1.0 (0.66, 1.0)	0.95 (0.82, 0.99)	96% (0.87)
Luciferase reporter phage assays (In-house) performed on culture isolates											
Banaiee (2003)[26]	Mexico	In-house LRP (luminometry)	BACTEC 460	Isolate	Yes	NR	7/65	0	1.0 (0.59, 1.0)	1.0 (0.95, 1.0)	100% (1.0)
Banaiee (2001)[27]	Mexico	In-house LRP (luminometry)	BACTEC 460	Isolate	Yes	NR	3/47	0	1.0 (0.29, 1.0)	1.0 (0.93, 1.0)	100% (1.0)
Hazbon (2003)[32]	Colombia	In-house LRP (luminometry)	Proportion method	Isolate	No	No	11/37	6	1.0 (0.72, 1.0)	0.89 (0.75, 0.97)	92% (0.79)
Hazbon (2003)[32]	Colombia	In-house LRP (Bronx box)	Proportion method	Isolate	No	No	10/34	14	1.0 (0.69, 1.0)	0.94 (0.80, 0.99)	95% (0.88)
Riska (1999)[39]	U.S.A	In-house LRP (Bronx box)	BACTEC 460	Isolate	No	No	10/17	0	1.0 (0.69, 1.0)	0.94 (0.71, 0.99)	96% (0.92)
Lu (2000)[35]	China	In-house LRP (luminometry)	Absolute concentration method	Isolate	Yes	Yes	13/9	0	0.92 (0.64, 0.99)	0.89 (0.52, 0.99)	91% (0.81)
Banaiee (2004)[41]	South Africa	In-house LRP (luminometry)	BACTEC 460	Isolate	Yes	Yes	9/182	4	1.0 (0.66, 1.0)	1.0 (0.98, 1.0)	100% (1.0)

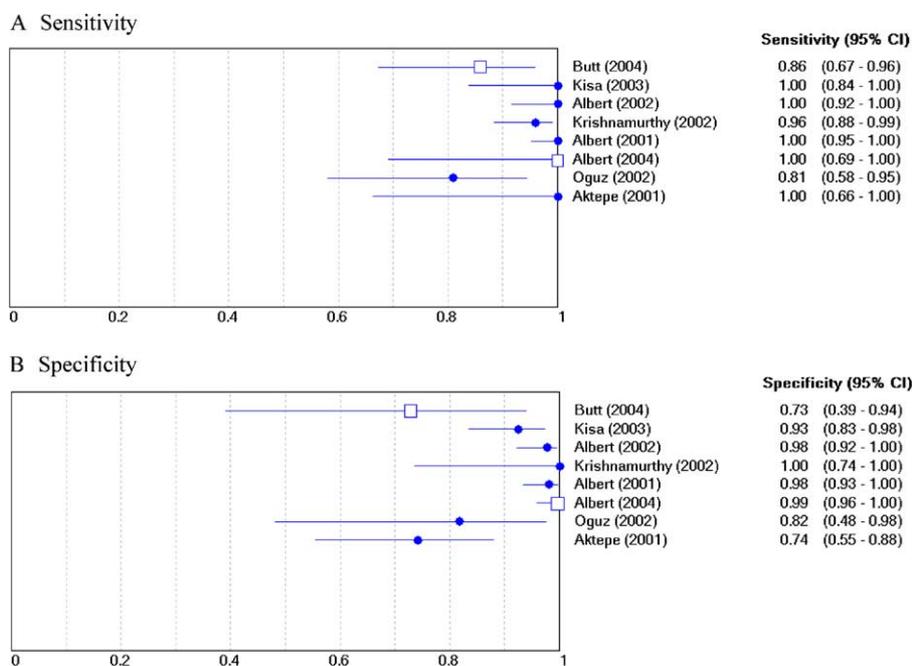LRP, luciferase reporter phage; NR, not reported; CI, confidence interval.

Figure 3   Forest plot of sensitivity and specificity for commercial phage amplification assays (FASTPlaque-TB tests) performed on culture isolates (*N*=6 studies), and clinical specimens (*N*=2 studies). (A) Sensitivity. (B) Specificity. The point estimates of sensitivity and specificity from each study are shown as solid circles (when performed on culture isolates) and squares (when directly performed on sputum specimens). Error bars are 95% confidence intervals (CI).

evaluated D-29 phage-based in-house assays. With the exception of two studies, all amplification assays used culture isolates. Two studies directly applied the FASTPlaque-TB assays on sputum specimens.[24,29]

Fig. 3 presents the forest plot of sensitivity and specificity estimates for the eight commercial phage amplification assays. The results are also displayed as a SROC curve in Fig. 4(A). As seen in Fig. 3, the sensitivity estimates ranged from 0.81 to 1.0. Five of 8 (63%) studies had a perfect sensitivity of 100%. The specificity estimates, on the other hand, were more variable than sensitivity (range 0.73-1.0). Although 4 of 8 (50%) studies had specificity estimates >95%, there were three studies with relatively poor accuracy (specificity ranging from 0.73-0.82). The SROC curve (Fig. 4(A)) shows a high level of overall accuracy with an area of 0.99, and Q* of 0.95. Of the two studies that performed phage assays directly on sputum specimens, one study[24] reported sensitivity and specificity of 1.0 and 0.99, respectively, whereas another reported sensitivity of 0.86 and specificity of 0.73.[29]

Fig. 5 presents the forest plot for the six in-house phage amplification assays. Both sensitivity (range 0.97-1.0) and specificity (range 0.84-1.0) estimates were consistently high. The SROC curve (Fig. 4(B)) shows a high level of overall accuracy with an area of 0.99, and Q* of 0.98.

The Table 1 presents the data on agreement between conventional DST and phage amplification (commercial and in-house) assays. Agreement estimates were high—11 of 14 (79%) studies had agreement >90%. The kappa estimates were also high, and ranged from 0.56 to 1.0.
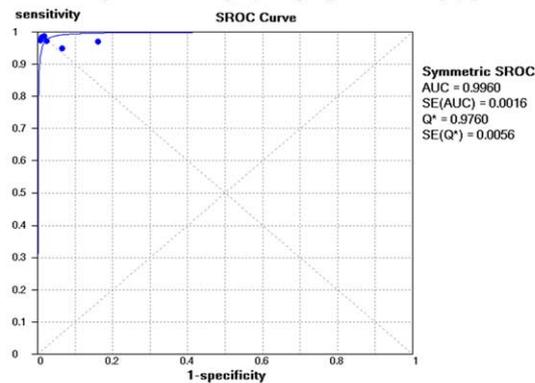
Accuracy of luciferase reporter phage assays

A total of seven studies evaluated phage assays that were based on luciferase reporter phages.[26,27,32,35,39,41] All of these were in-house assays, and all used culture isolates. None of the studies directly applied the LRP tests on sputum specimens. With the exception of two studies, all the LRP studies used luminometric detection of the light signal. Two studies used the Bronx box (a box containing Polaroid film) to detect light emission.[32,39]

Fig. 6 presents the forest plot of sensitivity and specificity estimates for the seven LRP assays. The results are also displayed as a SROC curve in Fig. 4(C). As seen in Fig. 6, all studies except one had a sensitivity of 100%. Specificity estimates, on the other hand, were slightly lower and more variable (range 0.89-1.0). The SROC plot (Fig. 4(C)) suggests high accuracy with an area of 0.98 and Q* of 0.95.

The Table 1 presents the data on agreement between conventional DST and LRP assays.

A Commercial amplification assays (*FASTPlaque-TB*) [N=8 studies]
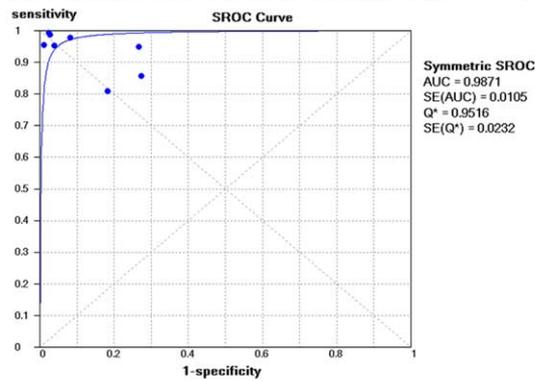
B In-house amplification assays (D29 phage-based assays) [N=6 studies]
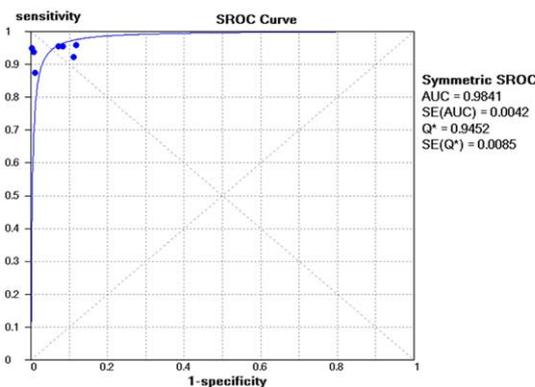

C Luciferase reporter phage assays [N=7 studies]

Figure 4   Summary ROC plots for phage-based assays, stratified by type of assay. (A) Commercial amplification assays (FASTPlaque-TB) (*N*=8 studies). (B) In-house amplification assays (D29 phage-based assays) (*N*=6 studies). (C) Luciferase reporter phage assays (*N*=7 studies). Each solid circle represents each study in the meta-analysis. The curve is the regression line that summarizes the overall diagnostic accuracy. SROC, summary receiver operating characteristic; AUC, area under the curve; SE (AUC), standard error of AUC; Q*, an index defined by the point on the SROC curve where the sensitivity and specificity are equal, which is the point closest to the top-left corner of the ROC space; SE(Q*), standard error of Q* index.

Agreement estimates were exceptionally high—all seven studies had agreement >90%. The kappa estimates were ranged from 0.79 to 1.0.

Impact of assay methodology on accuracy

Overall, the SROC plots stratified by type of assay (Fig. 4(A)–(C)) were fairly similar for all three assays (commercial and in-house phage amplification, and LRP assays), although commercial assays appear to produce lower and more variable accuracy estimates. The area under the SROC curve and Q* estimates were fairly similar, indicating comparable diagnostic accuracy.

Isoniazid resistance among rifampicin-resistant isolates

Seven studies provided data on INH resistance by conventional DST methods among strains that were resistant to rifampicin.[12,24,25,29,31,37,40] These studies showed that, on average, 96% of the rifampicin-resistant isolates were also resistant to INH. Rifampicin resistance, therefore, was a fairly good surrogate marker of MDR-TB in the populations studied.

Discussion

Principal findings and clinical implications

Our meta-analysis of 21 studies (with all but two studies using culture isolates) on phage-based assays for the detection of rifampicin resistance suggests that phage assays are associated with fairly high sensitivity and specificity, when applied to culture isolates. These results indicate that, as a group, phage assays show promise for the detection of rifampicin resistance in culture isolates when compared to conventional agar-based drug susceptibility tests and rapid methods such as BACTEC 460. Once an isolate was drawn from a growing culture, the average turnaround time for phage-based assays was 48–72 h, whereas with conventional DST, the turnaround time might vary from 1–2 weeks for BACTEC and MGIT, and 3–4 weeks for solid media-based methods. In addition to rapidity, phage-based assays have the advantage of being less expensive and technologically simpler than molecular (e.g. PCR) and liquid-media based tests (e.g. MGIT, BACTEC). This offers some advantages in resource-limited settings.

Although these results suggest that phage-based assays are fairly sensitive and specific for rifampicin
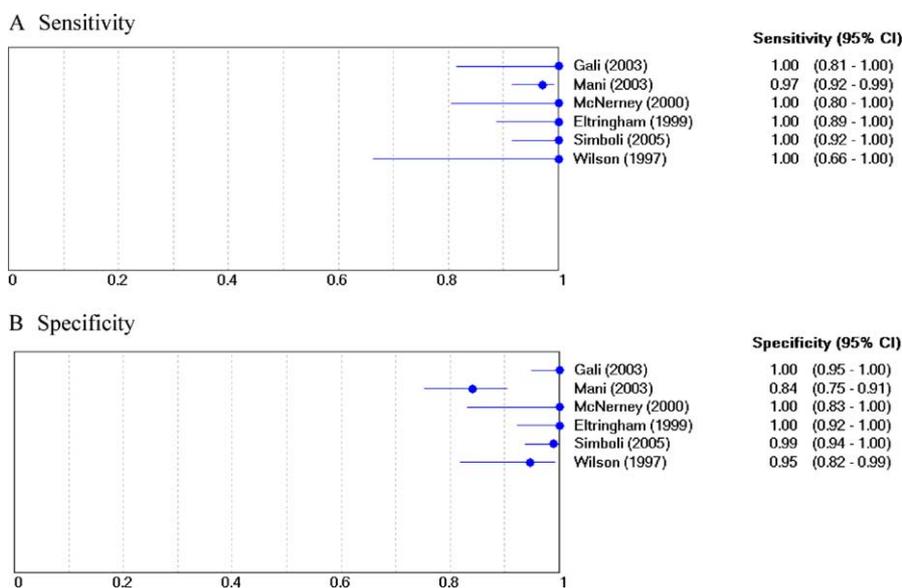
Figure 5 Forest plot of sensitivity and specificity for in-house phage amplification assays, all performed on culture isolates (N=6 studies). (A) Sensitivity. (B) Specificity. The point estimates of sensitivity and specificity from each study are shown as solid circles. Error bars are 95% confidence intervals (CI).

resistance, it is worth noting that phage-based assays do not show such high accuracy for the direct detection of *M. tuberculosis* in clinical specimens.[13] In a previous meta-analysis, we summarized data from 13 studies that evaluated commercial and in-house phage-based assays for the diagnosis of pulmonary TB.[13] The results suggested that phage-based assays have high specificity (range 0.83–1.00), but lower and variable sensitivity (range 0.21–0.88). This meta-analysis

also showed that phage-based assays have performance characteristics that are fairly similar to that of sputum microscopy—high specificity, but modest and variable sensitivity.[13] A major difference between the two meta-analyses is type of specimens used in the studies included; almost all studies (19 of 21) in the current meta-analysis used culture isolates, whereas all the studies (13 of 13) analysed in the previous meta-analysis directly performed the tests on sputum specimens. This
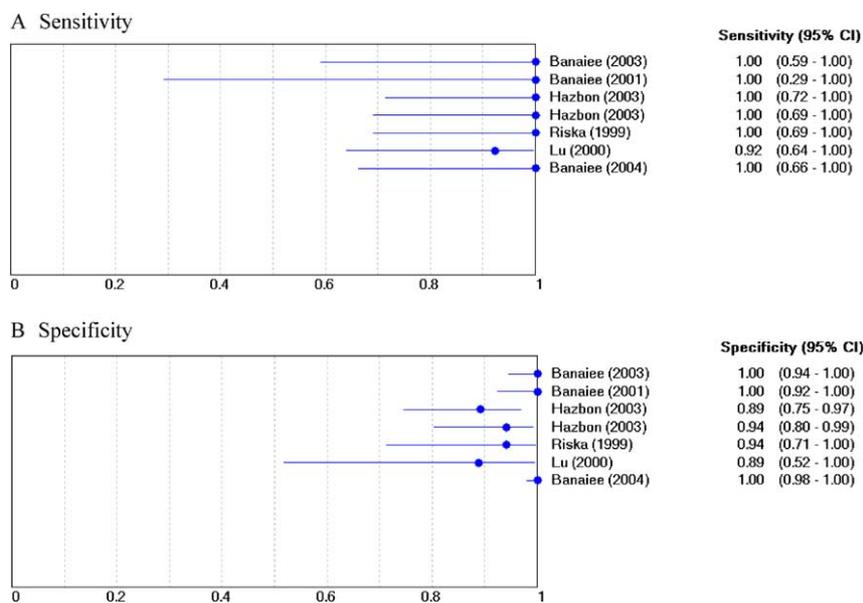

Figure 6 Forest plot of sensitivity and specificity for luciferase reporter phage assays, all performed on culture isolates (N=7 studies). (A) Sensitivity. (B) Specificity. The point estimates of sensitivity and specificity from each study are shown as solid circles. Error bars are 95% confidence intervals (CI).

difference in the specimens used has major implications for accuracy of phage-based assays and their clinical utility.

Although phage assays for rifampicin resistance are usually performed after primary isolation of *M. tuberculosis*, their reasonably high accuracy has greater clinical implications if they that can be directly applied to sputum specimens—a necessary condition in resource-limited settings where TB burden is greatest, and where cultures are not easy to obtain. Unfortunately, only two of 21 studies in our meta-analysis evaluated phage assays directly on sputum specimens.[24,29] Albert and colleagues evaluated the new FASTPlaque-TB-Response test directly on sputum specimens and demonstrated a sensitivity and specificity of 100 and 99%, respectively.[24] However, Butt and co-workers using the FASTPlaque-TB-RIF test showed a more modest sensitivity of 86% and specificity of 73%.[29] This difference in accuracy could be partly due to the variation in definitions used for rifampicin resistance and cut-points used for valid results. In the study by Albert and colleagues, assay results were considered valid if ≥ 100 plaques were observed on the RIF$-$ plate, and a strain was determined to be resistant if ≥ 50 plaques were seen on the RIF$+$ sample.[24] In the study by Butt and colleagues, results were considered valid if ≥ 20 plaques were seen on the RIF$-$ plate, and a strain was determined to be resistant when ≥ 20 plaques were seen on RIF$+$ plates.[29] This variability in cut-points emphasizes the need for better standardization of phage assays, and for further research to identify optimal thresholds for validity and test-positivity.

Our results suggest that more research is needed to establish the accuracy of phage-based assays when clinical specimens are used for testing. If they are found to be highly accurate for sputum specimens, they have the potential to make a significant clinical impact, particularly by reducing the time to diagnosis. They can be used to identify patients that are most likely to have MDR-TB, fail treatment or not responding to conventional anti-tuberculosis treatment. They can also be used to individualize drug regimens to suit the resistance pattern of the organism. This can be very helpful in MDR-TB 'hot spots' where the prevalence of drug resistance is very high.[1]

However, before phage assays can be successfully used in routine practice, several concerns have to be addressed. These include unexplained low sensitivity and specificity in some studies, and potential for contamination and indeterminate/uninterpretable results. These concerns have been raised previously in the context of phage assays for TB diagnosis.[13,42] They will also be important in the context of DST. Some studies reported low specificity estimates (in the range of 80%), and this is cause for concern.[22,36,38] In settings with a relatively lower prevalence of MDR-TB, false positive rates in the range of 10-20% can result in significant false over-reporting of MDR-TB. Because bacteriophages can replicate in non-tuberculous mycobacteria (NTM) as well as *M. tuberculosis*, there is always a potential for false positive results when phage assays are directly applied to sputum specimens. Some non-tuberculous mycobacteria (e.g. *M. smegmatis*) are naturally resistant to rifampicin, and in populations with high background prevalence of NTM, there is a risk of over-diagnosing MDR-TB. Because of the toxicity associated with MDR-TB treatment, it is important to study this issue further and determine the causes of false positive results in various settings. As noted earlier, rifampicin resistance may not be a perfect surrogate marker of MDR-TB in all settings.[2] In such populations, phage assay results for rifampicin, if used in isolation, might result in over-diagnosis of MDR-TB. For all these reasons, to minimize false positive results, a second confirmatory test (e.g. conventional DST for first line drugs) may be necessary to confirm and validate all positive phage results. However, this facility may not be easily available in resource-limited settings.

Strengths and limitations of the review

Our systematic review had several strengths. An important strength of our study was its comprehensive search strategy. We searched several databases and sources, and also identified studies by contacting authors, experts and test manufacturers. We attempted to limit publication bias by including journal articles as well as conference abstracts, and by including non-English language articles. Screening, study selection, and quality assessment were done independently and reproducibly by two reviewers. We minimized the problem of missing data by contacting authors. A majority of the authors responded by sharing additional information on their primary studies. We performed meta-analyses and exploration of heterogeneity in accordance with published guidelines.[14,17,18] To account for variation in assay methods, we performed subgroup analyses and analysed the phage amplification assays separately from the LRP assays, and also examined tests performed on cultures separately from those used on clinical specimens.

Our review had some limitations. Firstly, we did not address issues such as cost-effectiveness,

reliability, the incremental benefit of adding phage assays to other conventional tests, and the net effect of phage assays on clinical care and patient outcomes. None of the studies in our meta-analysis reported data on the clinical impact of detection of rifampicin resistance on long-term treatment outcomes such as cure rates, treatment failures and mortality. Another problem concerns the exclusion of phage assays that produced indeterminate and/or contaminated results, and intermediate sensitivity from the analyses in the primary studies. Such exclusions might have overestimated the sensitivity and specificity estimates in our meta-analysis. Lastly, despite our comprehensive literature searches, it is still possible that the high accuracy of phage-based assays in our meta-analysis could be due to publication bias. Although statistical (e.g. Begg and Egger tests) and graphical methods (e.g. funnel plots) are available to detect potential publication bias in meta-analyses of randomized controlled trials, such techniques have not been adequately evaluated for diagnostic data.[17] It is, therefore, difficult to rule out publication bias in our meta-analysis.

Implications for research

Future studies on phage assays must include the following design features: consecutive or random selection of patients with suspected drug-resistant TB, blinded interpretation of both phage and conventional DST results, and an adequate number of patients with confirmed rifampicin resistance. Studies should explicitly report the cut-points used to determine test validity and drug resistance, concentrations of the drugs used, and the proportion of phage tests that are indeterminate or contaminated. If indeterminate or contaminated results are excluded from the final analysis, the impact of such exclusions on accuracy measures should be reported. In addition, more studies are needed on the accuracy of phage assays when directly applied to sputum specimens. The challenge in such studies would be to develop methods to increase the sensitivity of phage-based assays without compromising specificity. Another challenge would be to extend the scope of phage assays to cover other first-line anti-tuberculosis drugs. This is already being attempted with some success.[8,9] Lastly, studies should provide data on prevalence of non-tuberculous mycobacteria and the effect of NTM isolation on phage assay specificity.

Studies are also needed to establish the cost-effectiveness and operational, logistical (e.g. turn-around time), and clinical advantages of phage assays over conventional DST. The advantage in getting quicker results is clearly greater with direct testing of sputum specimens, as compared to performing the phage assays after primary culture isolation. Although only amplification-based assays are currently available as commercial kits, it is necessary to also simplify and develop LRP based assays into a standardized, low-technology kit that could be made available commercially. This is currently being attempted on a small scale.[32,39,41] Lastly, studies must evaluate the actual impact of detection of rifampicin resistance on patient care and clinical decision making. It will be necessary to show that testing for rifampicin resistance using phage assays can improve treatment outcomes, reduce morbidity and mortality among TB patients, and decrease transmission of MDR-TB.

Conclusions

Current evidence is largely restricted to the use of phage assays for the detection of rifampicin resistance in culture isolates. When applied to isolates, these assays appear to have relatively high sensitivity and specificity. Specificity estimates, however, appear to be relatively lower and more variable than sensitivity, and this can potentially result in over-diagnosis of MDR-TB in low prevalence settings. In contrast, evidence is lacking on the accuracy of these assays when they are directly applied to sputum specimens. If phage-based assays can be directly used on clinical specimens and if they are shown to have high accuracy, they have the potential to improve the diagnosis and management of MDR-TB. However, before phage assays can be successfully used in routine practice, several concerns have to be addressed, including unexplained low sensitivity and specificity in some studies, potential for contamination and indeterminate results, and variability in cut-points used to determine test validity and drug resistance. Better standardization of the assay procedure, therefore, is needed for widespread applicability.

Acknowledgements

References

1. Espinal MA. The global situation of MDR-TB. *Tuberculosis (Edinb)* 2003;**83**(1-3):44-51.

2. World Health Organization. *Anti-tuberculosis drug resistance in the world. Third global report*. Geneva: World Health Organization; 2004.

3. Fisher M. Diagnosis of MDR-TB: a developing world problem on a developed world budget. *Expert Rev Mol Diagn* 2002; **2**(2):151-9.

4. Heifets LB, Cangelosi GA. Drug susceptibility testing of *Mycobacterium tuberculosis*: a neglected problem at the turn of the century. *Int J Tuberc Lung Dis* 1999;**3**(7):564-81.

5. Parsons LM, Somoskovi A, Urbanczik R, Salfinger M. Laboratory diagnostic aspects of drug resistant tuberculosis. *Front Biosci* 2004;**9**:2086-105.

6. Hazbon MH. Recent advances in molecular methods for early diagnosis of tuberculosis and drug-resistant tuberculosis. *Biomedical* 2004;**24**(Suppl 1):149-62.

7. Somoskovi A, Parsons LM, Salfinger M. The molecular basis of resistance to isoniazid, rifampin, and pyrazinamide in *Mycobacterium tuberculosis*. *Respir Res* 2001;**2**(3):164-8.

8. Mole RJ, Maskell T. Phage as a diagnostic—the use of phage in TB diagnosis. *J Chem Technol Biotechnol* 2001;**76**(7):683-8.

9. Trollip A, Albert H, Maskell T. Bacteriophage-based technologies for the rapid diagnosis and drug susceptibility testing of tuberculosis. *Am Clin Lab* 2001;**20**(9):39-42.

10. Jacobs Jr WR, Barletta RG, Udani R, Chan J, Kalkut G, Sosne G, et al. Rapid assessment of drug susceptibilities of *Mycobacterium tuberculosis* by means of luciferase reporter phages. *Science* 1993;**260**(5109):819-22.

11. McNerney R. TB: the return of the phage. A review of fifty years of mycobacteriophage research. *Int J Tuberc Lung Dis* 1999;**3**(3):179-84.

12. Wilson SM, Al-Suwaidi Z, McNerney R, Porter J, Drobniewski F. Evaluation of a new rapid bacteriophage-based method for the drug susceptibility testing of *Mycobacterium tuberculosis*. *Nat Med* 1997;**3**(4):465-8.

13. Kalantri SP. Bacteriophage-based tests for the detection of *Mycobacterium tuberculosis* in clinical specimens: a systematic review and meta-analysis. Master's in Public Health (MPH) Thesis, University of California, Berkeley; 2005.

14. Pai M, McCulloch M, Enanoria W, Colford Jr JM. Systematic reviews of diagnostic test evaluations: what's behind the scenes? *ACP J Club* 2004;**141**(1):A11-A13.

15. Pai M, McCulloch M, Gorman JD, Pai N, Enanoria W, Kennedy G, et al. Systematic reviews and meta-analyses: an illustrated, step-by-step guide. *Natl Med J India* 2004; **17**(2):86-95.

16. Hadgu A. Discrepant analysis: a biased and an unscientific method for estimating test sensitivity and specificity. *J Clin Epidemiol* 1999;**52**(12):1231-7.

17. Deeks JJ. Systematic reviews in health care: systematic reviews of evaluations of diagnostic and screening tests. *BMJ* 2001;**323**(7305):157-62.

18. Irwig L, Tosteson AN, Gatsonis C, Lau J, Colditz G, Chalmers TC, et al. Guidelines for meta-analyses evaluating diagnostic tests. *Ann Intern Med* 1994;**120**(8):667-76.

19. Littenberg B, Moses LE. Estimating diagnostic accuracy from multiple conflicting reports: a new meta-analytic method. *Med Decis Making* 1993;**13**(4):313-21.

20. Walter SD. Properties of the summary receiver operating characteristic (SROC) curve for diagnostic test data. *Stat Med* 2002;**21**(9):1237-56.

21. Lijmer JG, Bossuyt PM, Heisterkamp SH. Exploring sources of heterogeneity in systematic reviews of diagnostic tests. *Stat Med* 2002;**21**:1525-37.

22. Aktepe OC, Altindis M, Esen N. *A rapid method for rifampicin susceptibility testing of Mycobacterium tuberculosis 11th European Congress of Clinical Microbiology and Infectious Diseases (ECCMID), 2001, Istanbul, Turkey*.

23. Albert H, Heydenrych A, Mole R, Trollip A, Blumberg L. Evaluation of FASTPlaque TB-RIF, a rapid, manual test for the determination of rifampicin resistance from *Mycobacterium tuberculosis* cultures. *Int J Tuberc Lung Dis* 2001; **5**(10):906-11.

24. Albert H, Trollip A, Seaman T, Mole RJ. Simple, phage-based (FASTPplaque) technology to determine rifampicin resistance of *Mycobacterium tuberculosis* directly from sputum. *Int J Tuberc Lung Dis* 2004;**8**(9):1114-9.

25. Albert H, Trollip AP, Mole RJ, Hatch SJB, Blumberg L. Rapid indication of multidrug-resistant tuberculosis from liquid cultures using FASTPlaqueTB-RIF, a manual phage-based test. *Int J Tuberc Lung Dis* 2002;**6**(6):523-8.

26. Banaiee N, Bobadilla-del-Valle M, Riska PF, Bardarov S, Small PM, Ponce-de-Leon A, et al. Rapid identification and susceptibility testing of *Mycobacterium tuberculosis* from MGIT cultures with luciferase reporter mycobacteriophages. *J Med Microbiol* 2003;**52**(Pt 7):557-61.

27. Banaiee N, Bodadilla-del-Valle M, Bardarov S, Riska PF, Small PM, Ponce-de-Leon A, et al. Luciferase reporter mycobacteriophages for detection, identification, and antibiotic susceptibility testing of *Mycobacterium tuberculosis* in Mexico. *J Clin Microbiol* 2001;**39**(11):3883-8.

28. Bardarov S, Dou H, Eisenach K, Banaiee N, Ya S, Chan J, et al. Detection and drug-susceptibility testing of M-tuberculosis from sputum samples using luciferase reporter phage: comparison with the mycobacteria growth indicator tube (MGIT) system. *Diagn Microbiol Infect Dis* 2003;**45**(1):53-61.

29. Butt T, Ahmad RN, Afzal RK, Mahmood A, Anwar M. Rapid detection of rifampicin susceptibility of *Mycobacterium tuberculosis* in sputum specimens by mycobacteriophage assay. *J Pak Med Assoc* 2004;**54**(7):379-82.

30. Eltringham IJ, Drobniewski FA, Mangan JA, Butcher PD, Wilson SM. Evaluation of reverse transcription-PCR and a bacteriophage-based assay far rapid phenotypic detection of rifampin resistance in clinical isolates of *Mycobacterium tuberculosis*. *J Clin Microbiol* 1999;**37**(11):3524-7.

31. Gali N, Dominguez J, Blanco S, Prat C, Quesada AD, Matas L, et al. Utility of an in-house mycobacteriophage-based assay for rapid detection of rifampin resistance in *Mycobacterium tuberculosis* clinical isolates. *J Clin Microbiol* 2003;**41**(6): 2647-9.

32. Hazbon MH, Guarin N, Ferro BE, Rodriguez AL, Labrada LA, Tovar R, et al. Photographic and luminometric detection of luciferase reporter phages for drug susceptibility testing of clinical *Mycobacterium tuberculosis* isolates. *J Clin Microbiol* 2003;**41**(10):4865-9.

33. Kisa O, Albay A, Bedir O, Baylan O, Doganci L. Evaluation of FASTPlaquetb-RIF for determination of rifampicin resistance in *Mycobacterium tuberculosis* complex isolates. *Int J Tuberc Lung Dis* 2003;**7**(3):284–8.

34. Krishnamurthy A, Rodrigues C, Mehta AP. Rapid detection of rifampicin resistance in *M. tuberculosis* by phage assay. *Ind J Med Microbiol* 2002;**20**(4):211–4.

35. Lu B, Fu Z, Xu S. Rapid rifampicin susceptibility test by using recombinant mycobacteriophages. *Zhonghua Jie He He Hu Xi Za Zhi* 2000;**23**(8):480–4.

36. Mani C, Selvakumar N, Kumar V, Narayanan S, Narayanan PR. Comparison of DNA sequencing, PCR-SSCP and PhaB assays with indirect sensitivity testing for detection of rifampicin resistance in *Mycobacterium tuberculosis*. *Int J Tuberc Lung Dis* 2003;**7**(7):652–9.

37. McNerney R, Kiepiela P, Bishop KS, Nye PM, Stoker NG. Rapid screening of *Mycobacterium tuberculosis* for susceptibility to rifampicin and streptomycin. *Int J Tuberc Lung Dis* 2000; **4**(1):69–75.

38. Oguz VA, Guneri S, Erdenizmenli M, Yapar N, Kuruuzum Z, Cakmak R, et al. Investigation of rifampin resistance in *Mycobacterium tuberculosis* by two different methods. *Toraks Dergisi* 2002;**3**(2):173–7.

39. Riska PF, Su Y, Bardarov S, Freundlich L, Sarkis G, Hatfull G, et al. Rapid film-based determination of antibiotic susceptibilities of *Mycobacterium tuberculosis* strains by using a luciferase reporter phage and the Bronx box. *J Clin Microbiol* 1999;**37**(4):1144–9.

40. Simboli N, Takiff H, McNerney R, Lopez B, Martin A, Palomino JC, et al. In-house phage amplification assay is a sound alternative for detecting rifampin-resistant *Mycobacterium tuberculosis* in low-resource settings. *Antimicrob Agents Chemother* 2005;**49**(1):425–7.

41. Banaiee N, January V, Barthus C, Lambride M, RoDiti D, Behr MA, et al. *Semi-automation and evaluation of luciferase reporter mycobacteriophages for susceptibility testing of Mycobacterium tuberculosis* cultures in South Africa. *American Society for Microbiology 104th general meeting 2004, New Orleans, LA* 2004 [Abstract U-018].

42. Takiff H, Heifets L. In search of rapid diagnosis and drug-resistance detection tools: is the FASTPlaqueTB test the answer? *Int J Tuberc Lung Dis* 2002;**6**(7):560–1.